



Finding Order in the Machine

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):

Procter, R., & Johnson, C. (Ed.) (2002). Finding Order in the Machine. In C. Johnson (Ed.), *Proceedings of the European Annual Conference on Human Decision Making and Control*

Published in:

Proceedings of the European Annual Conference on Human Decision Making and Control

Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.



Finding Order in the Machine
Mark Hartswood¹, Rob Procter¹, Roger Slack¹, Mark Rouncefield²
¹ICCS/HCRC, Division of Informatics, University of Edinburgh
mjh|rnp|rslack@cogsci.ed.ac.uk
²Department of Computing, University of Lancaster
m.rouncefield@lancaster.ac.uk

Abstract: This paper examines the use of a computer-aided detection system for mammography and argues that it is important to appreciate work practice in order to understand the ways that the system will be used outwith the context of clinical trials. As an illustration of the lacuna, the ‘missing what’ in clinical trials, we show how order is found in the lived work of using the system and how readers make sense of its behaviour in the work setting. The paper concludes with a call for the use of ethnography in trials as a means of explicating the uses of technologies in real-world situations.

Keywords: mammography; ethnography; reading practice; clinical trials.

Introduction

The aim of this paper is to show how clinical trials elide the ‘lived work’ (Livingston, 1986) of, for example, doing reading mammograms. We show how such trials presumptions leave out the wider interactional contexts of doing work in real world settings; what we would call the sociality of work. The paper argues for the introduction of context-sensitive methodologies such as ethnography to explicate this lived work. Such a technique would consider technologies at the point of use in the work setting as opposed to the trial context. It is only through such considerations that we can appreciate concepts such as performance impact, usability and utility in their fullest sense.

Technology in Action

Breast cancer accounts for one fifth of all deaths from cancer among women in the UK. Established in 1988, the goal of the UK breast screening programme is to achieve a reliable cancer detection rate. Clinicians or readers¹ of mammograms, i.e., radiological images of the breast, are required to find what might be small and faint features in complex visual environments and to ensure that they detect as many cancers as possible (true positives) while keeping the number of women recalled unnecessarily (false positives) as low as possible. A number of computer-aided detection (CAD) systems have been developed which analyse the mammograms and ‘prompt’ readers to look at suspicious features; this paper details the trial of one such system.

The CAD system on trial consists of two components – the mammogram scanning and analysing unit and the mammogram viewing box with built-in displays for visual prompts. Mammograms on the viewing box are scrolled up and down. The prompts are synchronised with the mammograms, but the timing of their presentation is controlled by the reader. The system prompts for two types of features that are early indicators of breast cancer: micro-calcification clusters -- small deposits of calcium visible as tiny bright specks; ill-defined and stellate lesions -- areas of radiographically-dense tissue appearing as a bright patch that might indicate a developing tumour. The former are marked by a shaded triangle, the latter by an asterisk and a circle is drawn around either prompt type if the system’s confidence is high.

Readers were observed doing the various trial sets and then asked about their experiences of using the prompts. Readers were also taken back to cases identified in the trial set where they had appeared to have had difficulty or spent a long time making their decision, and asked to talk through any problems or issues to do with the prompts and their decisions. Although there were variations in how readers approached a reading and the trial, the fieldwork extract below gives some idea of the process observed:

Case 1: Gets blank mammogram to mask area of the mammogram (so I can concentrate on it ... these are set up differently from the way I usually look at them ... so I have to train my eye each time."). Using magnifying glass. Marking on booklet. Looking from booklet to scan. Homing in on an area -- "I'd say it's benign."

¹ Most (but not all) readers are qualified radiologists. We will use the more general term of reader.

Case 2: Using blank mammogram. Takes mammogram off roller and realigns. Magnifying glass. Looking from booklet to mammogram. "I'd not recall ... what the computer has picked up is benign ... it may even be talcum powder."

Case 10: Looking at mammogram - using blank mammogram to mask area. Magnifying glass. Looking at booklet prompts - looking back at mammogram. "This is a case where without the prompt I'd probably let it go ... but seeing the prompt I'll probably recall ... it doesn't look like a mass but she's got quite difficult dense breasts ... I'd probably recall."

As with everyday reading so with the trial, readers used a repertoire of manipulations to make certain features 'more visible'. A magnifying glass may be used to assess the shape, texture and arrangement of calcifications or, where the breast is dense, the mammogram may be removed and taken to a separate light box. Where a reader wished to attend to a particular segment of the mammogram, another mammogram may be used to blank off a part of it. In cases where a suspicious feature was seen on one view, readers used their fingers or an object such as a pen for measurement and calculation. These repertoires of manipulations are an integral part of the embodied practice of reading mammograms.

Strengths of the CAD system in supporting this kind of work lay in picking up subtle signs that some readers felt they might have missed and stimulating interaction between reader and the available technology by prompting them to re-examine the mammogram. Of course, this does not carry within it a decision as to what to do with the prompted feature – that requires decisions to be made by the readers. Readers also frequently express the opinion that they are better at 'spotting' some cancers -- as having skills or deficiencies in noticing particular types of feature within mammograms. This was another area where the CAD prompts were seen as useful, as both compensating in some (consistent) way for any individual weaknesses of the reader and as a reminder of good practice.

Two sources of uncertainty for readers can be found in deciding whether a mammogram is not recallable: first, a detected feature warrants a recall (if the feature is 'sufficiently suspicious' or regarded as artefactual and so on) and, second, satisfaction of search (when does one give up looking for features?). The aim of the CAD system is to deal with the second dimension as opposed to the first. Our previous studies have shown how readers reflexively adapt their work practices in order to build and sustain their 'professional vision' (Goodwin, 1994), and that this, in turn, contributes to the management of individual and collective performance. Readers have evolved an 'ecology of practice' for performance management that is deployed as part of the routine of the work. (Hartswood, Procter, Rouncefield and Slack, 2002). Through artful use of the public character of the screening reporting form and annotation work, readers use double reading to make their work observable-reportable to manage the uncertainties mentioned above -- when it is arguably most salient -- as they do it. Our interest here is to examine the role of the CAD system in the management of uncertainty: does the system manage uncertainty or create yet more areas of uncertainty for readers?

The CAD system should not be taken to make things less uncertain – decisions still have to be made and these fall to the readers (c.f. case 10 above). The prompts are, so to speak, docile in that their character is simply to prompt, as opposed to say what should be done. In the above fieldwork, we see that readers attempt to ascertain what a prompted feature is. That a prompt occurs is a meaningful thing, but what to do about it is still a readers' matter. It seems to us that the missing elements in these trials – the readers' sense making is not taken into account. There is still a deal of sense-making to be done in order to account for what the system is showing as accountably this or that feature warranting this or that decision. In other words, the system still requires the professional vision of the reader to remedy prompts as what they accountably are.

The question is how readers make sense of the CAD system. Following Schütz, we might argue that readers render mammograms intelligible using a mosaic of 'recipe knowledge': "a kind of organisation by habits, rules and principles which we regularly apply with success." (Schütz, 1972:73). While the common experiences and rules embodied in the mosaic are always open to potential revision they are, nevertheless, generally relied upon for all practical purposes as furnishing criterion by which adequate sense may be assembled and practical activities realised. Unlike everyday interaction, the CAD system cannot repair what it 'means', and difficulties can arise as readers rush to premature and often mistaken conclusions about what has happened, what is happening, what the system 'meant', what it 'is thinking', and so on

(Hartswood and Procter, 2000). The reader using the system for the first time is confronted with docile prompts and what the meaning of these becomes apparent only when one looks at the natural history of what the system's prompts have come to (obviously this can only be done retrospectively).

Everyday Reading Work and the Problem of Clinical Trials

Our initial evaluation of the CAD system also raises a number of questions concerning the appropriateness of quasi-clinical trials for technological innovations. Divorced from the lived reality of everyday reading work and the various affordances of the work setting (such as the annotated record, previous mammograms or a medical biography), the value of such a trial for the deployment of the technology in a (very different) real setting is in some doubt. Following Berg (1997), we argue that new technologies must become part of local work practices and practical reasoning – they are used in context and not in some abstract ideal-typical setting, such as are supposed in clinical trials. The record of achievement in the field of clinical support systems is patchy: many systems perform well in laboratory trials but are found wanting in use. The design rationale for clinical support systems, for example, often assumes generic difficulties, whereas clinicians' needs may be highly specific. The clinicians' problem becomes re-formulated in terms of what the technology can do, rather than their actual needs.

In that they are closely linked to existing technology affordances, the work setting practical actions and practical reasoning of clinicians raises important questions for the design and use of new technologies. They also raise questions as to whether the changes in practice changes that new technologies are intended to support are actually achievable. We might, indeed, question why one seeks to replace current practice as opposed to supporting it: the use of new technologies seems too often to reconfigure as opposed to supporting existing practice. This image of medical technology as panacea seems to be borne out by the logic and conduct of trials: in such contexts it is possible to eliminate all 'worldly' contingencies such as work practices and thereby to suggest a clear perspective on the impact of the tool. We argue that things are not as clear-cut in that trials take place in a rarefied world – what happens out in the 'messy' 'real world' is a different matter.

Conclusions

Reader training programmes for CAD systems may need to be designed to provide not only a resource for initial familiarisation, but also to support the continued learning of users and evolving of practices. The issue of change over time also raises some wider issues of evaluation in that many of the benefits of new technology are unlikely to be evident until several years or more after its introduction and adaptation to the particular circumstances of use. Yet, inevitably, evaluations are set up to investigate evidence of immediate benefits. The experimental context in which these trials are undertaken elide the social character of the work involved and thereby erase some of the most crucial dimensions of readers' work – we would argue that these need to be put back in order to have technologies that afford work in real settings as opposed to clinical trials. We do not advocate scrapping clinical trials, rather the point is to put back some of the context and to explicate tool use in context: such an appreciation would provide a more robust investigation of what tools actually do in practical work settings.

References

- Berg, M. (1997). *Rationalising Medical Work: Decision Support techniques and Medical Practices*. Cambridge: MIT Press.
- Goodwin, C. (1994). *Professional Vision*. *American Anthropologist*. 96; 606-633.
- Hartswood, M. and Procter, R. (2000). Computer-aided mammography: a case study of coping with fallibility in a skilled decision-making task. *Journal of Topics in Health Information Management*, vol. 20(4), May, p. 38-54.
- Hartswood, M, Procter, R., Rouncefield, M. and Slack, R. (2002). *Performance Management in Breast Screening: A Case Study of Professional Vision and Ecologies of Practice*. *Journal of Cognition, Technology and Work*, vol. 4(2), p. 91-100.
- Livingston, E. (1986) *The Ethnomethodological Foundations of Mathematics* London: RKP.
- Schütz, A. (1972) *The Phenomenology of the Social World* London: Heinemann Educational.